Election manipulation: the average-case

Joint works with Elchanan Mossel and Ariel Procaccia

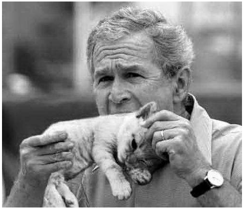
Miklós Z. Rácz

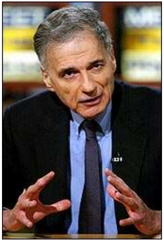
UC Berkeley

EconCS Seminar

Berkeley, April 9, 2013.

US Election 2000





Votes in Florida

48.84%	48.85%	1.64%

Nader supporters could have voted strategically and elected Gore.

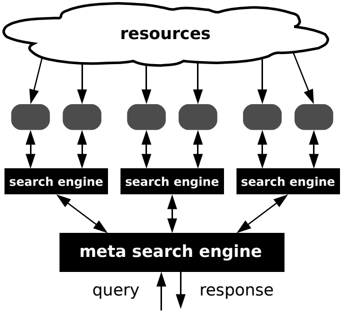
Artificial Intelligence & Computer Science

Virtual elections a standard tool in preference aggregation

- Elections can solve planning problems in multiagent systems (Ephrati and Rosenschein, 1991)
- Web metasearch engine (Dwork et al., 2001)
  - engines = voters, web pages = candidates

Threat of manipulation relevant, since software agents

- have computing power,
- have no moral obligation to act honestly.



Outline

Outline

Social Choice Theory

Quantitative Social Choice

Proof ideas

Coalitions

Social Choice Theory

- Social Choice Theory
  is the theory of collective decision making
- Originates from Condorcet's voting paradox,
  late 18th century
- Theory developed in Economics in 1950-70s
- Celebrated results are negative:
  - Arrow's impossibility theorem (1950):
    "irrationality" of ranking 3 or more candidates
  - Gibbard-Satterthwaite theorem (1973-75):
    any non-dictatorial way of electing a winner out of 3 or more candidates can be manipulated

Basic Setup

- n voters, k candidates
- Each voter ranks the candidates:
  vote of voter i denoted by $\sigma_i \in S_k$
- Social Choice Function (SCF)
  $f : S_k^n \to [k]$ selects a winner:

$$\sigma = (\sigma_1, \ldots, \sigma_n) \mapsto f(\sigma)$$

Manipulation by a single voter:

Definition

The SCF f is manipulable by voter i if there exist two ranking profiles $\sigma = (\sigma_i, \sigma_{-i})$ and $\sigma' = (\sigma_i', \sigma_{-i})$ such that

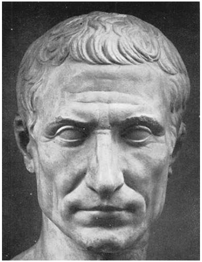
$$f(\sigma') \overset{\sigma_i}{>} f(\sigma).$$

Manipulability

- Ideal: nonmanipulable SCF.
- Q: when is this possible?

- Dictatorship:

$$\text{dict}_i(\sigma) := \text{top}(\sigma_i)$$



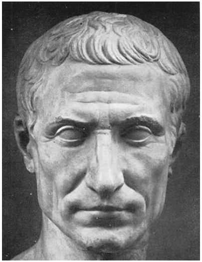
- ...anything socially acceptable?

Manipulability

- Ideal: nonmanipulable SCF.
- Q: when is this possible?

- Dictatorship:

$$\mathrm{dict}_i(\sigma) := \mathrm{top}(\sigma_i)$$



- ...anything socially acceptable?

Manipulability

- ...anything socially acceptable?

- For 2 candidates:
  strategyproofness is equivalent to monotonicity
- For 3 or more candidates: no such examples.

Theorem (Gibbard-Satterthwaite, 1973-75)

Every SCF that takes on at least three values and is not a dictator is manipulable.

Outline

Social Choice Theory

Quantitative Social Choice

Proof ideas

Coalitions

Is there a way around manipulation?

Two lines of research:

- Are there SCFs where it is *hard* to manipulate?
- Can manipulation be avoided with *good probability*?

Assumption: large number of voters and/or candidates.

Computational hardness of manipulation

Idea: election is vulnerable to manipulation only if it can be computed efficiently.

- Bartholdi, Tovey, Trick (1989): there exists a voting rule, such that it is NP-hard to compute a manipulative vote.
- Bartholdi, Orlin (1991): manipulation is NP-hard for Single Transferable Vote (Oakland mayor elections)
- ...many other developments...
- Problem: relies on NP-hardness as a measure of computational difficulty
- Is it hard *on average*? What if it is *typically* easy to manipulate?

Quantitative Social Choice

Basic question: is it possible to avoid manipulation with very good probability?

$\rightsquigarrow$ Random rankings

- Kelly, 1993: Consider people voting uniformly and independently at random; i.e. $\sigma \in S_k^n$ is uniform.
- Q: What is the probability of manipulation?

$$M(f) := \mathbb{P}(\sigma : \text{ some voter can manipulate } f \text{ at } \sigma)$$

- Gibbard-Satterthwaite theorem: If f takes on at least 3 values and is not a dictator, then

$$M(f) \geq \frac{1}{(k!)^n}$$

- If manipulation is so unlikely, perhaps we do not care?

Quantitative Social Choice

If f is "close" to a dictator $\rightsquigarrow M(f)$ can be very small
Quantifying distance:

$$\mathbf{D}(f,g) = \mathbb{P}(f(\sigma) \neq g(\sigma))$$
$$\mathbf{D}(f,G) = \min_{g \in G} \mathbb{P}(f(\sigma) \neq g(\sigma))$$

Assumption: f is ε-far from nonmanipulable functions:

$$\mathbf{D}(f, \mathrm{NONMANIP}) \geq \varepsilon$$

Conjecture (Friedgut, Kalai, Nisan (2008))

If $k \geq 3$ and $\mathbf{D}(f, \mathrm{NONMANIP}) \geq \varepsilon$, then

$$M(f) \geq poly\left(n, k, \varepsilon^{-1}\right)^{-1},$$

and a random manipulation works.
In particular: manipulation is easy on average.

Results

Theorem (Friedgut, Kalai, Keller, Nisan (2008,2011))

For $k = 3$ *candidates, if* $\mathbf{D}(f, \text{NONMANIP}) \geq \varepsilon$ *then*

$$M(f) \geq c\frac{\varepsilon^6}{n}.$$

If, in addition, f *is neutral, then*

$$M(f) \geq c'\frac{\varepsilon^2}{n}.$$

Neutrality of f: treats all candidates in the same way, i.e. is invariant under permutation of the candidates.

No computational consequences, since $k = 3$.

Note: some dependence on n is needed, see e.g. plurality: $O\left(n^{-1/2}\right)$ probability of manipulation.

Results, cont'd

Theorem (Isaksson, Kindler, Mossel (2010,2012))

If $k \geq 4$ and f is neutral, then $\mathbf{D}(f, \text{NONMANIP}) \geq \varepsilon$ implies

$$M(f) \geq poly\left(n, k, \varepsilon^{-1}\right)^{-1}.$$

Moreover, the trivial algorithm for manipulation works.

Computational consequences.

Removing neutrality:

Theorem (Mossel, R. (2012))

If $k \geq 3$ and $\mathbf{D}(f, \text{NONMANIP}) \geq \varepsilon$, then

$$M(f) \geq poly\left(n, k, \varepsilon^{-1}\right)^{-1}.$$

Moreover, the trivial algorithm for manipulation works.

Why is removing neutrality important?

- Anonymity vs. neutrality:
  - conflict, coming from tie-breaking rules
  - common SCFs anonymous $\rightsquigarrow$ not neutral
- In virtual election setting, neutrality can be not natural, e.g.:
  - (meta)search engine might treat websites in different languages in a different way
  - child-safe (meta)search engine:
    cannot have adult websites show up
- Sometimes candidates cannot be elected from the start
  - Local elections in Philadelphia, 2011
  - Dead man on NY State Senate 2010 election ballot
    (he received 828 votes)

Outline

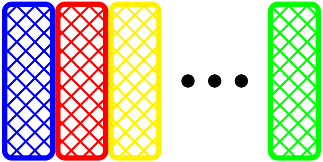
Social Choice Theory

Quantitative Social Choice

Proof ideas

Coalitions

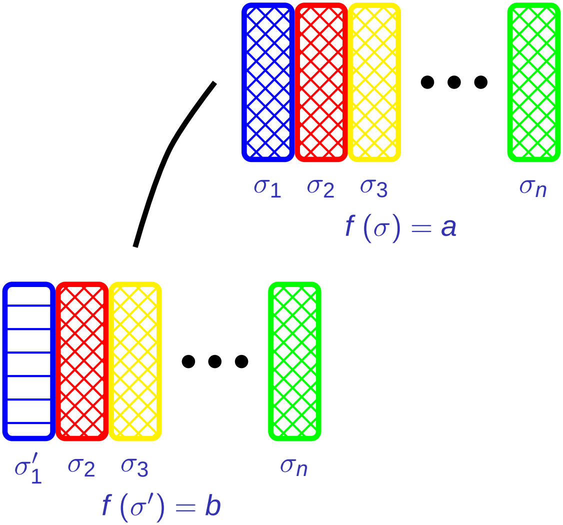
Finding manipulation points



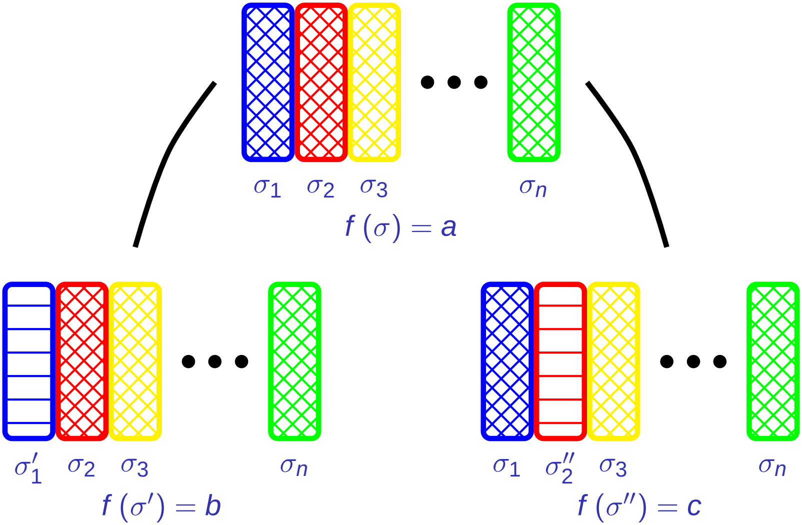
$\sigma_1 \quad \sigma_2 \quad \sigma_3 \qquad\qquad \sigma_n$

$$f(\sigma) = a$$

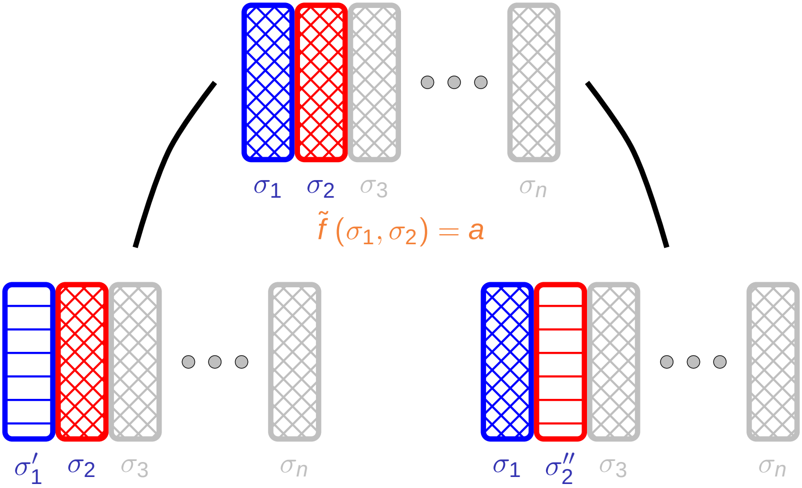
Finding manipulation points



Finding manipulation points



Finding manipulation points



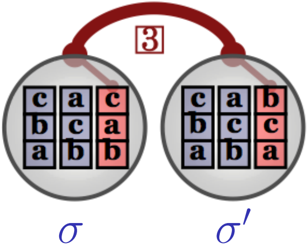
$$\tilde{f}\left(\sigma_1, \sigma_2\right) = a$$

$$\tilde{f}\left(\sigma_1', \sigma_2\right) = b \qquad \tilde{f}\left(\sigma_1, \sigma_2''\right) = c$$

$$\tilde{f}\left(\cdot, \cdot\right) := f\left(\cdot, \cdot, \sigma_{\{3,\ldots,n\}}\right)$$

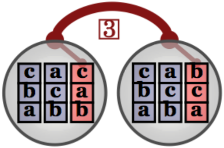
Rankings Graph

- Vertices: ranking profiles $\sigma \in S_k^n$
- Edges: if differ in one coordinate, i.e. (σ, σ') is an edge in voter i if $\sigma_j = \sigma'_j$ for all $j \neq i$, and $\sigma_i \neq \sigma'_i$



- SCF $f : S_k^n \to [k]$ induces a partition of the vertices
- Manipulation point can only occur on a boundary
- Boundary between candidates a and b in voter i: $B_i^{a,b}$.

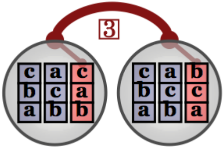
Boundary edges



σ σ'

$f(\sigma) = a$ $f(\sigma') = b$

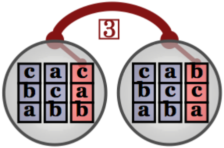
This edge is monotone and nonmanipulable.



σ σ'

$f(\sigma) = a$ $f(\sigma') = c$

This edge is monotone-neutral and manipulable.



σ σ'

$f(\sigma) = b$ $f(\sigma') = c$

This edge is anti-monotone and manipulable.

Isoperimetry

Recall: $k \geq 3$, uniform distribution, $\mathbf{D}\left(f, \mathrm{NONMANIP}\right) \geq \varepsilon$.

Lemma (Isoperimetric Lemma, IKM (2009))

There exist two voters $i \neq j$ such that $B_i^{a,b}$ and $B_j^{c,d}$ are big, i.e.

$$\mathbb{P}\left(\left(\sigma, \sigma^{(i)}\right) \in B_i^{a,b}\right) \geq \frac{\varepsilon}{poly\left(n,k\right)}, \quad \mathbb{P}\left(\left(\sigma, \sigma^{(j)}\right) \in B_j^{c,d}\right) \geq \frac{\varepsilon}{poly\left(n,k\right)},$$

where $c \notin \{a, b\}$.

If f is neutral, may assume $\{a, b\} \cap \{c, d\} = \emptyset \rightsquigarrow$ IKM (2009)

Now: assume $B_1^{a,b}$ and $B_2^{a,c}$ are big.

Fibers

- Partition the graph further, into so-called *fibers*
- Idea due to Friedgut, Kalai, Keller, Nisan (2008,2011)
- Ranking profile $\sigma \in S_k^n$ induces a vector of preferences between a and b:

$$x^{a,b} \equiv x^{a,b}(\sigma) = \left(x_1^{a,b}(\sigma), \ldots, x_n^{a,b}(\sigma)\right)$$

where $x_i^{a,b}(\sigma) = 1$ if $a \overset{\sigma_i}{>} b$, and $x_i^{a,b}(\sigma) = -1$ otherwise.

- A *fiber*: $F\left(z^{a,b}\right) := \left\{\sigma : x^{a,b}(\sigma) = z^{a,b}\right\}$
- Can partition the graph according to fibers:

$$S_k^n = \bigcup_{z^{a,b} \in \{-1,1\}^n} F\left(z^{a,b}\right)$$

Small and large fibers

Can also partition the boundaries according to the fibers:

$$B_1\left(z^{a,b}\right) := \left\{\sigma \in F\left(z^{a,b}\right) : f(\sigma) = a, \exists \sigma' \text{ s.t. } \left(\sigma, \sigma'\right) \in B_i^{a,b}\right\},$$

Distinguish between *small* and *large* fibers for boundary $B_1^{a,b}$:

Definition (Small and large fibers)

Fiber $B_1\left(z^{a,b}\right)$ is *large* if

$$\mathbb{P}\left(\sigma \in B_1\left(z^{a,b}\right) \middle| \sigma \in F\left(z^{a,b}\right)\right) \geq 1 - \text{poly}\left(n, k, \varepsilon^{-1}\right)^{-1},$$

and *small* otherwise.

Notation:

$\text{Lg}\left(B_1^{a,b}\right)$: union of large fibers for the boundary $B_1^{a,b}$

$\text{Sm}\left(B_1^{a,b}\right)$: union of small fibers for the boundary $B_1^{a,b}$

Cases

Recall: boundaries $B_1^{a,b}$ and $B_2^{a,c}$ are big.

Cases:

- $\mathrm{Sm}\left(B_1^{a,b}\right)$ is big
- $\mathrm{Sm}\left(B_2^{a,c}\right)$ is big
- $\mathrm{Lg}\left(B_1^{a,b}\right)$ and $\mathrm{Lg}\left(B_2^{a,c}\right)$ are both big

Large fiber case

Assume $\mathrm{Lg}\left(B_1^{a,b}\right)$ and $\mathrm{Lg}\left(B_2^{a,c}\right)$ are both big.

Two steps:

- Reverse hypercontractivity implies that the *intersection* of $\mathrm{Lg}\left(B_1^{a,b}\right)$ and $\mathrm{Lg}\left(B_2^{a,c}\right)$ is also big
- Gibbard-Satterthwaite implies that if $\sigma \in \mathrm{Lg}\left(B_1^{a,b}\right) \cap \mathrm{Lg}\left(B_2^{a,c}\right)$, then there exists manipulation point $\hat{\sigma}$ "nearby": σ and $\hat{\sigma}$ agree in all except perhaps the first two coordinates.

$\rightsquigarrow$ many manipulation points.

Small fiber case (sketch)

Assume $\mathrm{Sm}\left(B_1^{a,b}\right)$ is big.

1. By isoperimetric theory, for every small fiber $B_1\left(z^{a,b}\right)$, the size of the boundary, $\partial B_1\left(z^{a,b}\right)$, is comparable:

$$\left|\partial B_1\left(z^{a,b}\right)\right| \geq \mathrm{poly}\left(n, k, \varepsilon^{-1}\right)^{-1}\left|B_1\left(z^{a,b}\right)\right|$$

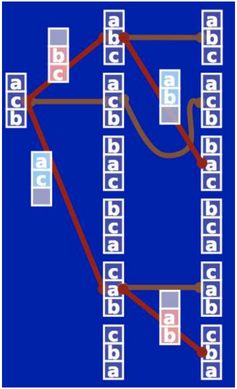
2. If $\sigma \in \partial B_1\left(z^{a,b}\right)$ in some direction $j \neq 1 \rightsquigarrow$ there exists a manipulation point $\hat{\sigma}$ "nearby", i.e. σ and $\hat{\sigma}$ agree in all but two coordinates
3. If $\sigma \in \partial B_1\left(z^{a,b}\right)$ in direction 1, then either there exists a manipulation point $\hat{\sigma}$ "nearby", or fixing coordinates 2 through n, we have a dictator on the first coordinate.
4. Look at the boundary of the set of dictators
$\rightsquigarrow$ manipulation point nearby.

Subtleties...

- We cheated in a few places...
- Most importantly, when we apply Gibbard-Satterthwaite, we lose a factor of $(k!)^2$...
- OK for constant number of candidates, but not for large k.

Refined rankings graph

- To get polynomial dependency, use refined rankings graph
- $(\sigma, \sigma') \in E$ if σ, σ' differ in a *single voter* and an *adjacent transposition*
- Need to prove: geometry = refined geometry, up to $\mathrm{poly}(k)$ factors.
- Need to prove:
  combinatorics still works
- Gives manipulation by permuting only a few adjacent candidates



Outline

Social Choice Theory

Quantitative Social Choice

Proof ideas

Coalitions

What if there is a coalition of voters?

- Various closely related “manipulation” problems
- Examples:
  - Coalitional manipulation
  - Bribery
- Various types of “manipulation”:
  - Constructive
  - Destructive
- Is the coalition specified?
  - yes: decision problem
  - no: optimization problem

Unifying framework

- Xia (2012) general results:
  - votes are i.i.d.,
  - SCF is a generalized scoring rule,
  - then w.h.p. the number of vote operations needed is either 0, $\Theta\left(\sqrt{n}\right)$, $\Theta\left(n\right)$, or ∞
- More specific results available for
  - specific problems,
  - specific voting rules,
  - specific distributions (e.g., uniform).

Bribery

- Question: can the winner be changed...
  - ...by any coalition of a particular size?
  - ...by a specific coalition of a particular size?
  - ...to a specific candidate? (constructive)
  - ...to just any other candidate? (destructive)
- Procaccia & Rosenschein (2007),Xia & Conitzer (2008):
  - votes are i.i.d.,
  - SCF is a generalized scoring rule,
  - if the coalition size is $o\left(\sqrt{n}\right)$, then w.h.p. all such coalitions are powerless
  - if the coalition size is $\omega\left(\sqrt{n}\right)$, then w.h.p. such a coalition is all-powerful
- Pritchard and Wilson (2009):
  - votes are uniform,
  - SCF is a scoring rule,
  - then minimum size of a succesful manipulating coalition is $C\left(w\right)\sqrt{n}$, where the distribution of $C\left(w\right)$ is explicit.

Smooth phase transition for bribery

- Question: can the winner be changed...
  - ...by any coalition of a particular size?
  - ...by a specific coalition of a particular size?
  - ...to a specific candidate? (constructive)
  - ...to just any other candidate? (destructive)

- Suppose the coalition has size $c\sqrt{n}$.
- Question: what is the phase transition like as c goes from 0 to ∞?
- Mossel, Procaccia, R. (2012): smooth when
  - votes are i.i.d. (and $p(\pi) \geq \delta$ and $\mathbb{P}(W_a) \geq \varepsilon$),
  - SCF is a generalized scoring rule.

Take aways

- Robust impossibility theorems: manipulation is computationally easy on average
- Interesting math involved

Take aways

- Robust impossibility theorems: manipulation is computationally easy on average
- Interesting math involved

Thank you!